Functional Annotation Report: *aceF* (PP_0338, UniProt Q88QZ6)

Acetyltransferase (E2) component of the pyruvate dehydrogenase complex

Organism: *Pseudomonas putida* KT2440 (strain ATCC 47054 / DSM 6125 / NCIMB 11950), a Gammaproteobacterium of the order Pseudomonadales.

---

1. Summary (Answer to the Research Question)

aceF encodes the **dihydrolipoyllysine-residue acetyltransferase (E2)** component of the **pyruvate dehydrogenase multienzyme complex (PDHc)** — **EC 2.3.1.12**. Its primary enzymatic function is to catalyze the transfer of an acetyl group from the reductively-acetylated lipoyl arm (S8-acetyldihydrolipoyl-lysine) to **coenzyme A**, producing **acetyl-CoA**:

> N6-[(R)-S8-acetyldihydrolipoyl]-L-lysyl-[protein] + CoA ⇌ N6-[(R)-dihydrolipoyl]-L-lysyl-[protein] + acetyl-CoA

Beyond catalysis, E2 is the **structural and organizational heart of PDHc**: 24 copies of its catalytic domain self-assemble into a hollow cube (octahedral 432 symmetry) that forms the core to which the peripheral E1 (pyruvate dehydrogenase, *aceE*/PP_0339) and E3 (dihydrolipoamide dehydrogenase, *lpd*) subunits attach. The enzyme functions in the **cytoplasm** and sits at the pivotal metabolic node linking sugar catabolism (in *P. putida*, principally the Entner–Doudoroff/EDEMP route converging on pyruvate) to the **TCA cycle** and acetyl-CoA-dependent biosynthesis.

Gene-identity verification: the symbol *aceF*, the protein description, the 2-oxoacid-dehydrogenase family assignment, and the InterPro domains (lipoyl/biotinyl-binding, 2-oxoacid DH acyltransferase, PSBD) **all mutually agree**. This is an unambiguous, well-characterized enzyme; no symbol-collision problem exists.

---

2. Molecular Identity and Domain Architecture

The 546-residue protein (UniProt Q88QZ6) shows the canonical modular architecture of a Gram-negative PDHc E2, confirmed from the UniProt feature table:

Region	Residues	Module	Role
Lipoyl domain 1	2–75	Biotinyl/lipoyl-binding (β-barrel)	Carries covalent (R)-lipoate on a conserved Lys (in the `…LESDKASMEIP…` motif); "swinging arm"
Lipoyl domain 2	117–191	Biotinyl/lipoyl-binding (β-barrel)	Second lipoyl-lysine (second `…LESDKASMEIP…` motif)
Linkers	Ala/Pro-rich	Flexible hinges	Allow the lipoyl arms to visit E1, E2 and E3 active sites
PSBD	245–282	Peripheral subunit-binding domain	Docks E1 and E3 onto the E2 core
Catalytic domain	~290–546	Acetyltransferase (chloramphenicol-acetyltransferase fold)	Acetyl transfer to CoA; core assembly

Notable points: - **Two lipoyl domains** — intermediate between *E. coli* (three lipoyl domains) and mammalian/*Bacillus* E2 (typically one to two). Multiple lipoyl domains increase the effective local concentration and reach of the acetyl-carrying arm. - The catalytic domain contains a conserved **His-Ser-Asn catalytic set**, all present in Q88QZ6 (see §3): catalytic **Ser468** (in the `SSLGH` motif), catalytic **His519** (in the `DHR` motif), and **Asn523** just downstream. Cofactor: covalently bound **(R)-lipoate** (UniProt COFACTOR) on **Lys41 and Lys157**.

Cofactor loading (activation): the lipoyl domains are catalytically inert until a lipoyl group is attached to their conserved lysines. This is done either *de novo* — **LipB** octanoylates the lysine and the radical-SAM enzyme **LipA** inserts two sulfur atoms to form lipoate — or by **LplA**-mediated salvage of exogenous lipoate; in the absence of this modification the dehydrogenase is inactive and aerobic metabolism is blocked (P 21209092). *P. putida* KT2440 encodes the orthologous lipoylation machinery.

3. Primary Catalytic Function and Mechanism

Reaction (EC 2.3.1.12): E2 catalyzes reversible **transacetylation** between the protein-bound dihydrolipoyl-lysine and CoA. In the physiological (PDHc) direction, E1 first decarboxylates pyruvate and reductively acetylates the E2 lipoyl-lysine; E2 then transfers that acetyl group to CoA to yield **acetyl-CoA**, leaving a reduced (dihydro)lipoyl arm.

Substrate/acyl specificity. aceF is an **acetyl-specific** transferase acting on the acetyl group derived from pyruvate. *P. putida* KT2440 encodes two other E2 acyltransferase paralogs with distinct acyl specificities — **sucB** (PP_4188, Q88FB0, EC 2.3.1.61, *succinyl*transferase of the 2-oxoglutarate dehydrogenase complex) and **bkdB** (PP_4403, Q88EQ0, branched-chain 2-oxoacid dehydrogenase E2). aceF is only **~33% identical** to each of these paralogs, yet **69% identical** to a true PDH-E2 ortholog (*A. vinelandii* E2p). This roughly two-fold difference confirms that aceF's acetyl specificity is orthology-defined and encoded in its divergent catalytic domain, functionally separating the pyruvate→acetyl-CoA node from the TCA-cycle 2-oxoglutarate step (sucB) and branched-chain amino-acid catabolism (bkdB).

Mechanistic role of the modular design (substrate channelling): 1. A **lipoyl domain** presents its lipoyl-lysine to the E1 active site, where it is reductively acetylated (S8-acetyldihydrolipoamide). 2. The **flexible Ala/Pro linkers** swing the acetylated arm into the **E2 catalytic channel**. Structural work on the near-identical *Azotobacter vinelandii* E2p core shows a **~29 Å active-site channel** in which **CoA enters from the inside of the cube and the lipoamide arm enters from the outside**, so the two substrates meet buried within the trimer interface
( ).
P 1549782
3. Acetyl transfer produces acetyl-CoA; the resulting **dihydrolipoyl arm** is then presented to E3, which reoxidizes it (regenerating oxidized lipoamide and reducing NAD+ via FAD).

This "swinging-arm" coupling channels reactive intermediates between three spatially separated active sites without releasing them to bulk solvent. Cryo-EM of the human complex shows that CoA binding **modulates the conformational landscape of the lipoyl domains**, indicating the arm dynamics are actively coupled to substrate occupancy
( ).
P 31130485

Catalytic residues (experimentally defined in the homolog; conserved in aceF): Site-directed mutagenesis with crystallography of *A. vinelandii* E2p (P 7703242) established the active-site chemistry: **His610** is the **general base** for proton transfer (His610→Cys reduced activity ~500-fold), **Ser558** provides **transition-state stabilization** (Ser558→Ala ~200-fold reduction), and **Asn614** activates proton transfer. All three are conserved in *P. putida* aceF at **His519 (DHR motif), Ser468 (SSLGH motif), and Asn523**. Notably, aceF retains the **rare Asn** at the 614-equivalent position — a feature described as "exceptional" in *A. vinelandii* (most E2 homologs have Asp there) — reflecting their close Pseudomonadales kinship and validating *A. vinelandii* E2p as the structural surrogate for aceF.

---

4. Structural / Scaffolding Role

E2 is described structurally and functionally as the **central enzyme** of PDHc. Key evidence, from high-resolution structures of the close Pseudomonadales relative *A. vinelandii* E2p:

- The catalytic domain **forms a 24-subunit oligomer with octahedral 432 symmetry** (P 8487300); UniProt independently annotates Q88QZ6 as forming "a 24-polypeptide structural core with octahedral symmetry."
- **Eight tightly-associated trimers assemble into a hollow truncated cube** (~120–125 Å edge) with pores on each face, "forming the core of the multienzyme complex" (P 1549782). The trimer is the true building block; two levels of contacts (3-fold trimer + 2-fold trimer–trimer) build the cube.
- Each catalytic subunit has a **topology identical to chloramphenicol acetyltransferase (CAT)**, the structural basis for its acyl-transfer chemistry (P 1549782).

The **PSBD** provides the attachment platform: in the *E. coli* PDHc (the best-studied Gram-negative model), point substitutions in the PSBD (R129E, R150E) severely reduce complex activity and disrupt binding of both E1 and E3 as well as reductive acetylation of E2

( ).
P 23580650
Thus E2 both **builds the core** and **recruits the peripheral catalytic subunits**, giving PDHc its megadalton multienzyme architecture.

Quantitative homology (validation of the structural inference): a full-length global (Needleman–Wunsch) alignment of aceF (546 aa) gives **69.2% identity (444/642) to *A. vinelandii* E2p** (P10802, the crystallized/mutationally-dissected model) and **49.6% (316/637) to *E. coli* AceF** (P06959). This exceptionally high identity to a same-order (Pseudomonadales) enzyme means its solved structures and mechanism transfer to aceF with high confidence; length differences arise mainly from lipoyl-domain copy number and Ala/Pro linker length rather than the conserved catalytic core.

Complex partners in KT2440: aceF (E2, PP_0338, Q88QZ6) assembles with **E1** = **aceE (PP_0339, Q88QZ5, 881 aa)** and the shared **E3** = **lpdG (PP_4187, Q88FB1, dihydrolipoyl dehydrogenase, 478 aa)**. The adjacent loci PP_0338/PP_0339 are consistent with a co-transcribed **aceEF operon**, while the distal E3 (lpdG) is typical of a dihydrolipoyl dehydrogenase shared among the 2-oxoacid dehydrogenases and glycine-cleavage system.

---

5. Localization

- **Cytoplasm** (GO:0005737, and as a component of the cytoplasmic PDH complex, GO:0045254). As a bacterial enzyme, it operates in the cytosol — there is no mitochondrion; this is the prokaryotic counterpart of the mitochondrial-matrix mammalian PDHc-E2.

---

6. Pathway Context and Physiological Role

- **Metabolic node:** PDHc performs the irreversible oxidative decarboxylation of pyruvate → acetyl-CoA + CO2 + NADH, the principal gateway from central sugar catabolism into the TCA cycle and into acetyl-CoA-dependent pathways (fatty-acid synthesis, PHA/ polyhydroxyalkanoate biosynthesis, acetylation).
- **In *P. putida* KT2440 specifically:** glucose is catabolized predominantly through the **Entner–Doudoroff pathway** (KT2440 lacks a functional Embden–Meyerhof–Parnas glycolysis for net catabolism), with the cyclic **EDEMP** arrangement providing NADPH;

carbon converges on pyruvate, and PDHc (E1 *aceE*/PP_0339 – E2 *aceF*/PP_0338 – E3 *lpd*) feeds acetyl-CoA to the TCA cycle. The adjacent PP_0338/PP_0339 loci are consistent with an *ace* gene cluster encoding the E1 and E2 components.

- **Family:** 2-oxoacid dehydrogenase family. The E2 module is the paradigm shared with the 2-oxoglutarate (E2o, succinyltransferase EC 2.3.1.61) and branched-chain 2-oxoacid dehydrogenase complexes; *aceF* is the pyruvate-specific (acetyltransferase) member.

---

7. Evidence Summary

Claim	Evidence type	Source
EC 2.3.1.12; acetyl-transfer reaction; (R)-lipoate cofactor	Curated annotation (RuleBase/ UniRule)	UniProt Q88QZ6
Two lipoyl domains + PSBD + catalytic domain	Sequence/domain features	UniProt Q88QZ6; InterPro IPR000089, IPR003016, IPR001078, IPR006256
24-mer octahedral cubic core; CAT fold; 29 Å active-site channel	X-ray crystallography of Pseudomonadales homolog *A. vinelandii* E2p (2.6 Å)	P 1549782; P 8487300
Catalytic His610/Ser558/ Asn614 (→ aceF His519/ Ser468/Asn523)	Site-directed mutagenesis + crystallography (*A. vinelandii*)	P 7703242
Lipoyl domains require LipB/ LipA (de novo) or LplA (salvage) lipoylation	Biochemistry / proteomics (*E. coli*)	P 21209092
Lipoyl "swinging arm"; covalent lipoyl-lysine for active-site coupling	Biochemistry / MS mapping	P 21798751
PSBD tethers E1/E3 and enables reductive acetylation	Mutagenesis + structural MS (*E. coli*)	P 23580650
CoA-modulated lipoyl-domain dynamics / channelling	Cryo-EM + native MS (human PDHc)	P 31130485
Cytoplasmic localization; PDHc membership	Curated GO	UniProt Q88QZ6 (GO:0005737, GO:0045254, GO:0004742, GO:0006086)
69.2% identity to A. vinelandii E2p; 49.6% to E. coli AceF	Global sequence alignment (this work)	UniProt P10802, P06959
Partners aceE (PP_0339/ Q88QZ5) & lpdG (PP_4187/ Q88FB1); aceEF operon	Genomic loci / UniProt	UniProt Q88QZ5, Q88FB1

Claim	Evidence type	Source
Acetyl-specific: only ~33% identity to paralogs sucB (succinyl) & bkdB	Comparative alignment (this work)	UniProt Q88FB0, Q88EQ0

Strength of inference: No *P. putida*-specific enzymological study of aceF was located; however, the function is established at high confidence by (i) unambiguous, mutually-consistent UniProt/InterPro annotation, and (ii) direct structural/biochemical characterization of very close bacterial homologs (*A. vinelandii*, same order; *E. coli*, same class), which are the standard models for this enzyme family.

---

8. Supported vs. Refuted Hypotheses

Supported: - H1 — *aceF* is the E2 acetyltransferase (EC 2.3.1.12) of PDHc catalyzing acetyl-CoA formation. ✔ - H2 — *aceF* forms the 24-mer cubic core and scaffolds the complex via its PSBD. ✔ - H3 — Function depends on covalent lipoyl "swinging arms" enabling substrate channelling. ✔ - H4 — The enzyme acts in the cytoplasm at the glycolysis/ED–TCA junction. ✔ - H5 — aceF is acetyl-specific, distinct from the succinyl (sucB) and branched-chain (bkdB) E2 paralogs (~33% identity vs 69% to a PDH-E2 ortholog). ✔ - H6 — Catalytic His519/Ser468/Asn523 are conserved from the mutationally-validated *A. vinelandii* active site. ✔ (inferred)

Refuted / not applicable: - The gene-symbol-ambiguity contingency was ruled out: *aceF* unambiguously matches the annotated PDHc-E2 identity.

---

9. Limitations and Future Directions

- Direct experimental characterization is from homologs, not *P. putida* KT2440 itself; a KT2440-specific structure or kinetic study would confirm exact catalytic-residue geometry and the functional consequence of the two-lipoyl-domain arrangement.
- The precise lipoyl-lysine positions and the catalytic His were inferred from sequence motifs and homology; site-directed mutagenesis in KT2440 would confirm them.
- Whether both lipoyl domains are equally lipoylated/functionally redundant in KT2440 is untested.

- Regulatory features (e.g., any PDH-kinase control) and operon structure in KT2440 warrant genomic/proteomic confirmation.

Report generated during autonomous functional-annotation investigation (Iterations 1–5). Conclusions rest on curated UniProt/InterPro annotation, high-resolution structures and site-directed mutagenesis of close Pseudomonadales/Gammaproteobacteria homologs (A. vinelandii E2p, E. coli AceF), conserved-residue and quantitative sequence-identity analysis, and KT2440 genomic context.
